The Temporal Realization of Morphological and Syntactic Boundaries*

Ilse Lehiste

*Sponsored in part by the National Science Foundation through Grant GN-534.1 from the Office of Science Information Service to the Computer and Information Science Research Center, The Ohio State University.

The Temporal Realization of Morphological and Syntactic Boundaries

Ilse Lehiste

Abstract

This paper is concerned with the effect of morphological and syntactic boundaries on the temporal structure of spoken utterances. Two speakers produced twenty tokens each of four sets of words consisting of a monosyllabic base form, disyllabic and trisyllabic words derived from the base by the addition of suffixes, and three short sentences in which the base form was followed by a syntactic boundary, this in turn followed by a stressed syllable, one unstressed syllable, and two unstressed syllables. The sentences thus reproduced the syllabic sequences of the derived words. The duration of words and segments was measured from oscillograms. The manifestation of morphological and syntactic boundaries is discussed, and some implications of the findings relative to the temporal programming of spoken utterances are considered.

0. Introduction

This paper is concerned with the effect of morphological and syntactic boundaries on the temporal structure of spoken utterances. The investigation was prompted by the observation made in the course of a previous study, [1,2] that the duration of a word may be considerably reduced, if a derivational suffix is added to the word constituting the base. In this earlier study, the words <u>stead</u>, <u>skid</u> and <u>skit</u> were compared with <u>steady</u>, <u>skiddy</u> and <u>skitty</u>. It might have been expected that the latter set would be longer than the former by the average duration of the derivational suffix. It turned out instead that the duration of the base part of the derived word was considerably shortened, so that even with the addition of a fairly long -<u>y</u>, the overall duration of the derived words was not much different from that of the base words.

In the current study, four sets of words were examined, built around the base forms <u>stick</u>, <u>sleep</u>, <u>shade</u>, and <u>speed</u>. Each of the words occurred by itself and in eight additional utterance types. Five derivational suffixes were used, three of them monosyllabic and two disyllabic. The words were further placed in short sentences in which they were followed by a major syntactic boundary--the boundary between the noun phrase functioning as subject and the verb phrase functioning as predicate. The verb phrase itself either consisted of a stressed monosyllable (in three cases) or started with a stressed syllable (in one case); or it started with one or two unstressed syllables. The sentences thus reproduced the syllabic sequences of the derived words. It was the purpose of the study to explore whether there are any differences in the durations of the base, depending on whether it is followed by a morpheme boundary within the same word, or by a major syntactic boundary coinciding with the word boundary.

I. Method

The test material, presented in Table 1, was recorded by two speakers, R.G. (male) and L.S. (female), both graduate students at The Ohio State University. The recordings were made under standard conditions in an anechoic chamber using reliable recording equipment. The utterances were produced in two ways, to test the comparability of different contexts and to vary the fairly artificial recording technique of repeating the same word a large number of times. One of the ways was indeed the repetition technique: each word was uttered ten times under a subjectively established 'constant' rate. Then each set, consisting of base word, derived words, and three short sentences, was read ten times in succession. Each speaker thus produced 20 tokens of each word, for a total of 720 utterances by each speaker.

The durations of words and segments were measured from oscillograms, produced by processing the recorded tapes through a Frøkjær-Jensen Trans-Pitch Meter and Intensity Meter, connected to a four-channel Elema-Schönander Mingograph. The material was analyzed statistically, using the IBM 360 Model 75 computer available at The Ohio State University Instruction and Research Computer Center.

II. Comparability of the Two Sets of Data

For both sets of data, the following computations were carried through: the mean duration of each segment; the mean duration of each word; the mean duration of the base component of the derived word (e.g., stick in sticky); the variances and standard deviations of each segment and word. The differences between the corresponding means for each segment and word were tested for significance according to the formula:

$$(1) \qquad Z = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma^2_A}{N_A} + \frac{\sigma^2_B}{N_B}}}$$

The difference in variability between the two sets was tested by two (related) measures:[3]

$$(2) \qquad H = \frac{\sigma^2_{MAX}}{\sigma^2_{MIN}} \qquad\qquad C = \frac{\sigma^2_{MAX}}{\sigma^2_{MIN} + \sigma^2_{MAX}}$$

For the given number of tokens, the critical values (at the 95% confidence level) were 1.960 for Z, 4.030 for H, and 0.801 for C.

It was found that the differences between the two sets of utterances for each speaker were random, and that there was minimal overlap between the two speakers in cases of statistically significant differences. Out of 196 pairwise comparisons of $\bar{X}_A$ and $\bar{X}_B$, speaker R.G. had 65 significant differences, speaker L.S. 88 significant differences; the same segments were involved in 35 instances, but these segments constituted no natural set: there was no discernible system. A separate check of syllable nuclei showed 11 instances for R.G. and 26 instances for L.S. in which the means differed significantly, i.e. Z was higher than the critical value. The same syllable nucleus was involved in 9 instances. As regards the differences in variability between the two sets, speaker R.G. had 15 (out of 196) cases in which the difference in variances between the two sets was significant; speaker L.S. had 36 instances, of which 9 involved the same segment for both speakers. As far as syllable nuclei were concerned, speaker R.G. had 2 instances of significant differences, L.S. 4, with an overlap of 2.

Combining the two sets would tend to increase the extreme ranges for each combined set of utterances and thus increase the variability; but since the difference in variability between the two sets was negligible, it was decided to combine the two sets in future calculations. The resultant increase in variability was in effect quite small. It is hoped that the method of producing the test utterances in the two different ways described above will have reduced the artificiality of the situation in which long lists of words are produced out of context, and that the results are better applicable to a more natural speech situation.

III. Effect of Morpheme Boundaries

In order to study the effect of morpheme boundaries (and word boundaries) on the duration of the base to which derivative suffixes were added, B/D ratios were computed. This term refers to the ratio of the durations of the base word (produced by itself) and the sum of the durations of the same segments occurring in the derived word (e.g., the mean duration of stick would be divided by the mean duration of the stick part of the word sticky). These ratios were calculated for all test words, and, separately, for the syllable nuclei in all test words. The differences between the means were highly significant in all instances; Z-values, which were always higher than the critical value, will not be included in the tables. The results are presented in Tables II-V and graphically in Figures 1 - 4. The tables are self-explanatory; a few words of explanation may be needed for the figures.

On each figure, the derived word types and sentence types are given on the vertical axis. The horizontal axis is calibrated to show increasing B/D ratios. Points representing B/D ratios for words are connected with solid lines; points representing B/D ratios for syllable nuclei are connected with dashed lines. The curves start in the left hand top corner at the B/D value 1: Base/Base yields a ratio of 1. Increasing ratios show decrease in the duration of the base component of the derived word resp. its syllable nucleus.

Several observations may be made regarding the figures. In no case was the duration of the same set of segments greater in a derived word than in the base form. The suffixes -y, -er and -ing seem to be equivalent with respect to their effect on the duration of the stem. It appears that the number of segments in the suffix has no systematic effect on the duration of the stem. This observation is confirmed by looking at the behavior of stem forms before the suffix -ily. This suffix was in fact pronounced with a syllabic /l/ by both speakers in all productions; thus the stems of words like sticking and stickily were followed by two segments each, but the -ing suffix was monosyllabic and the -ily suffix was disyllabic. In all cases, the disyllabic -ily suffix produced greater reduction in the duration of the stem than the monosyllabic suffix -ing, although both consisted of the same number of segments.

The suffix -iness constitutes a special case. In each instance, the B/D ratio was greatest under this condition. This is a disyllabic suffix, as is -ily; however, its rhythmic structure is considerably different. It seems possible that in the case of the -iness suffix we are dealing with two cycles of derivation: that, for example, sticky is derived from stick in the first cycle, and stickiness from sticky in the second cycle. If this is so, then the ratios of stick /sticky and sticky/stickiness (involving the base forms stick and sticky respectively) should be approximately equal. Some support for this assumption may indeed be found in Table VI, which presents the pertinent ratios.

A comparison of the curves for words with the curves for syllable nuclei indicates that the reduction in the duration of a stem in the derived form is achieved more at the expense of vowels than at the expense of consonants. The nature of the vowel and the postvocalic consonant seem to play an equally important role. Intrinsically long syllable nuclei (like those in sleep, speed, and shade) are more compressible than intrinsically short syllable nuclei (as in stick). But /i/ in sleep, when followed by a voiceless plosive, is much less compressible than /i/ in speed and /eᴵ/ in shade. Tendencies for being reduced under a certain condition become accentuated when one looks at the most compressible segment: for both speakers, the greatest effects of the various positions are manifested in the syllable nuclei of speed and shade.

IV. Effect of Syntactic Boundaries

One of the hypotheses tested in this experiment was the hypothesis that syntactic boundaries would have temporal effects that are clearly distinct from those of morpheme boundaries. However, the results of this study show that as far as the temporal structure of utterances is concerned, effects of morpheme boundaries and effects of syntactic boundaries cannot be separated from each other. Furthermore, it is not certain that the boundaries as such have any effect at all, since the temporal structure of the utterances seems to depend most of all on their syllabic structure, regardless of the nature of the boundaries involved.

In sentences like Speed kills, we find durations of the test word that are very similar to those of disyllabic bimorphemic words;

sentences like The speed increased resemble most words like speediness, with an unstressed short syllable followed by a relatively long syllable. The addition of another unstressed syllable may have a further reducing effect, but the data are not consistent at this point. The major result here is the absence of any clear differences between the effects of morpheme boundaries and syntactic boundaries, and the likelihood that the durational structure is conditioned by the number of syllables rather than either by the number of segments or by the presence of boundaries.

V. Generality of the Findings

One of the ways to test the results would be to form predictions on the basis of these data and then compare the predictions with further observations. I intend to record other sets of words by the same speakers as well as the same sets of words by different speakers, and calculate the goodness of fit between predicted and observed B/D ratios. The basis for predictions might be Table VII, which combines words that seem to behave in a similar fashion for the two speakers.

VI. Discussion

The results of this study confirm earlier studies in some respects, but differ from them in certain important aspects.

Bolinger[4] stated that long syllables tend to acquire extra length if followed by another long syllable (long syllables being those that contain a full vowel); if followed by a short syllable, long syllables cannot acquire that extra length and therefore appear shorter. This process tends to ignore morpheme and word boundaries, and may take place across a syntactic boundary.

The present study confirms Bolinger's notion that temporal readjustment processes tend to ignore morpheme and word boundaries. The shortening of a long syllable before a short syllable is likewise confirmed in all the data. However, in sentences of the type Speed kills, the word speed (and words in analogous sentences) certainly did not acquire any extra length, at least in comparison to isolated productions of the same word.

Gaitenby[5] found a common ratio of segment-to-utterance length for all dialects of American English sampled in her study. When segment durations were converted to percentages of total utterance time, it was found that 90% of all the segments varied less than 5.3% for any speaker. The longer the utterance in terms of number of segments, the shorter the absolute duration of any given segment, until an approximate minimum duration was reached beyond which segments could not be compressed any further. She noted also that words immediately preceding a pause tended to expand in utterances of all lengths. According to Gaitenby, it would thus be the word closest to the pause that would acquire extra length, while in longer utterances, the preceding parts of the sentence would be produced at a faster rate. This seems to be borne out by the findings: in the three sentences, the base word became successively shorter, the farther it was removed from the end of the sentence. A difference between Gaitenby's results and those obtained in this study is the observation that utterance length should be determined with reference to number and type of syllables rather than with reference to the number of segments.

Chomsky and Halle[6] have postulated a hierarchy of boundaries which delimit linguistic units that serve as domains of application of different kinds of phonological rules. Although the authors are careful to state that phonetic effects need not be associated with (word) boundaries, the postulation of a hierarchy of boundaries naturally prompts a phonetician to look for possibly hierarchical differences in the manifestations of these boundaries. I had previously formulated the hypothesis that phonological units are definable in terms of suprasegmental patterns, while their boundaries are mainly manifested in terms of modifications of segments.[7] Few, if any, indications of word boundaries emerged from the present study. There were a small number of instances in which the duration of the segment preceding a word boundary was greater than the duration of the same segment preceding a morpheme boundary. As far as the overall temporal organization of the utterances is concerned, no evidence for a hierarchical organization of boundaries was found as a result of this study. The temporal organization of spoken language seems to take place in terms of speech production units which are fairly independent of the morphological or syntactic structure of the utterances.

Acknowledgements

Grateful acknowledgement is made of the help of Mr. Thomas G. Whitney of the Ohio State University Instruction and Research Computer Center, who wrote the computer programs employed in this study.

Footnotes

[1] I. Lehiste, "The Temporal Organization of Higher-Level Linguistic Units," Paper presented at the April 1970 meeting of the Acoustical Society of America, Atlantic City, N.J. (1970).

[2] I. Lehiste, "Temporal Organization of Spoken Language," In: Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen, Ed. by L. L. Hammerich, Roman Jakobson, and Eberhard Zwirner (Akademisk Forlag, Copenhagen, 1971), pp. 275-285.

[3] B. J. Winer, Statistical Principles in Experimental Design (McGraw-Hill, New York, 1962), p. 94.

[4] D. Bolinger, "Length, Vowel, Juncture," Linguistics 1.1.5-29 (1963). (To be revised).

[5] J. Gaitenby, "The Elastic Word," Paper given at the Tenth Annual National Conference on Linguistics, sponsored by the Linguistic Circle of New York, 13 March 1965. Status Report on Speech Research SR-2 (Haskins Laboratories, New York, 1965), pp. 3.1-3.12.

[6] N. Chomsky and M. Halle, The Sound Pattern of English (Harper & Row, New York, 1968).

[7] I. Lehiste, Suprasegmentals (M.I.T. Press, Cambridge, Mass., 1970).

Table I. Test materials used in the study. The symbol - is used to indicate the boundary between stem and derivative suffix. # symbolizes word boundary; ´ and ˘ refer to stressed and unstressed syllables.

BASE	stick	sleep	shade	speed
-Y	sticky	sleepy	shady	speedy
-ER	sticker	sleeper	shader	speeder
-ING	sticking	sleeping	shading	speeding
-ILY	stickily	sleepily	shadily	speedily
-INESS	stickiness	sleepiness	shadiness	speediness
# ´	the stick fell	sleep heals	the shade lingered	speed kills
# ˘ ´	the stick is broken	sleep refreshes	the shade increased	the speed increased
# ˘ ˘ ´	the stick was discarded	my sleep was disturbed	the shade was refreshing	the speed was controlled

Table II. Mean durations (in milliseconds), standard deviations and B/D ratios for two sets of words and corresponding syllable nuclei produced by speaker R.G.

Utterance	Duration of base	σ	B/D ratio	Duration of Syl. nucleus	σ	B/D ratio
stick	401.55	29.45		130.70	6.94	
sticky	312.80	23.68	1.284	93.45	6.53	1.399
sticker	302.50	17.49	1.327	89.45	8.85	1.461
sticking	295.45	16.92	1.359	88.80	7.28	1.472
stickily	291.10	17.90	1.379	84.15	6.75	1.553
stickiness	265.75	15.79	1.511	78.90	5.63	1.657
The stick fell	274.85	14.10	1.461	87.90	7.02	1.487
The stick is broken	248.20	12.65	1.618	81.65	7.57	1.601
The stick was discarded	245.10	13.49	1.638	77.90	5.81	1.678
sleep	409.80	18.96		123.55	14.55	
sleepy	336.80	19.70	1.217	84.15	7.97	1.468
sleeper	341.25	19.83	1.201	83.10	9.21	1.487
sleeping	330.35	18.12	1.241	81.50	10.11	1.516
sleepily	313.35	13.99	1.308	69.60	8.58	1.775
sleepiness	287.05	13.81	1.428	62.05	6.79	1.991
sleep heals	305.95	16.33	1.339	75.95	8.41	1.627
sleep refreshes	299.60	19.90	1.368	61.85	4.67	1.998
My sleep was disturbed	307.45	17.44	1.333	59.65	9.65	2.071

Table III. Mean durations (in milliseconds), standard deviations and B/D ratios for two sets of words and corresponding syllable nuclei produced by speaker L.S.

Utterance	Duration of base	σ	B/D ratio	Duration of Syl. nucleus	σ	B/D ratio
stick	431.80	43.33		168.90	23.25	
sticky	346.00	34.44	1.248	115.50	15.83	1.462
sticker	331.95	25.88	1.301	109.65	14.75	1.540
sticking	348.30	30.56	1.240	109.20	17.36	1.547
stickily	303.10	17.93	1.425	77.05	6.89	2.192
stickiness	271.60	20.78	1.590	76.50	6.92	2.208
The stick fell	311.15	22.74	1.388	91.35	11.17	1.849
The stick is broken	283.90	19.46	1.521	88.85	10.83	1.901
The stick was discarded	268.15	28.40	1.610	80.75	8.42	2.092
sleep	442.45	39.62		180.30	16.85	
sleepy	363.40	19.64	1.218	131.45	9.24	1.372
sleeper	363.35	22.87	1.218	127.25	8.90	1.417
sleeping	374.45	18.26	1.182	132.45	10.87	1.361
sleepily	342.60	16.72	1.291	114.50	8.72	1.575
sleepiness	307.70	16.39	1.438	96.55	8.45	1.867
Sleep heals	325.00	25.33	1.361	113.55	14.77	1.588
Sleep refreshes	282.75	18.96	1.565	93.55	9.74	1.927
My sleep was disturbed	314.90	26.82	1.405	99.40	19.27	1.814

Table IV. Mean durations (in milliseconds), standard deviations and B/D ratios for two sets of words and corresponding syllable nuclei produced by speaker R.G.

Utterance	Duration of base	σ	B/D ratio	Duration of Syl. nucleus	σ	B/D ratio
speed	511.50	34.95		266.00	28.17	
speedy	359.75	15.09	1.422	150.50	10.25	1.767
speeder	344.75	16.42	1.484	141.50	11.01	1.880
speeding	342.50	13.13	1.493	136.00	9.81	1.956
speedily	322.50	18.03	1.586	120.00	8.27	2.217
speediness	313.25	16.57	1.633	115.50	7.76	2.303
Speed kills	344.00	17.06	1.487	125.50	8.87	2.120
The speed increased	301.25	15.12	1.698	110.00	7.61	2.418
The speed was controlled	293.50	20.53	1.743	104.00	8.97	2.558
shade	454.10	28.88		266.15	18.61	
shady	327.20	20.08	1.388	181.85	14.79	1.464
shader	324.20	18.81	1.401	172.40	9.54	1.544
shading	306.95	23.39	1.479	158.00	11.24	1.684
shadily	276.70	10.20	1.641	132.05	8.74	2.016
shadiness	265.20	17.60	1.712	125.35	9.83	2.123
The shade lingered	324.80	18.49	1.398	146.95	16.23	1.811
The shade increased	298.60	18.44	1.521	130.15	12.93	2.045
The shade was refreshing	307.60	26.05	1.476	131.50	18.61	2.024

Table V. Mean durations (in milliseconds), standard deviations and B/D ratios for two sets of words and corresponding syllable nuclei produced by speaker L.S.

Utterance	Duration of base	σ	B/D ratio	Duration of Syl. nucleus	σ	B/D ratio
speed	574.25	30.00		297.85	16.25	
speedy	394.85	23.89	1.454	163.30	11.69	1.824
speeder	403.85	18.44	1.422	171.75	13.52	1.734
speeding	396.10	24.54	1.450	158.75	12.86	1.876
speedily	354.50	29.75	1.620	126.25	16.98	2.359
speediness	322.70	23.41	1.780	104.40	6.66	2.853
Speed kills	416.55	27.28	1.379	163.05	19.07	1.827
The speed increased	342.85	20.97	1.675	127.30	11.68	2.340
The speed was controlled	305.50	22.00	1.880	96.65	7.92	3.082
shade	454.65			267.70	22.88	
shady	321.65	20.72	1.413	165.25	11.26	1.620
shader	326.75	26.61	1.391	160.50	14.16	1.668
shading	312.95	22.09	1.453	159.30	19.72	1.680
shadily	294.15	26.41	1.546	139.95	21.89	1.913
shadiness	261.65	25.37	1.738	112.55	11.65	2.378
The shade lingered	331.95	36.24	1.370	154.40	23.93	1.734
The shade increased	282.20	22.03	1.611	135.75	16.90	1.972
The shade was refreshing	273.40	23.97	1.633	114.25	15.97	2.343

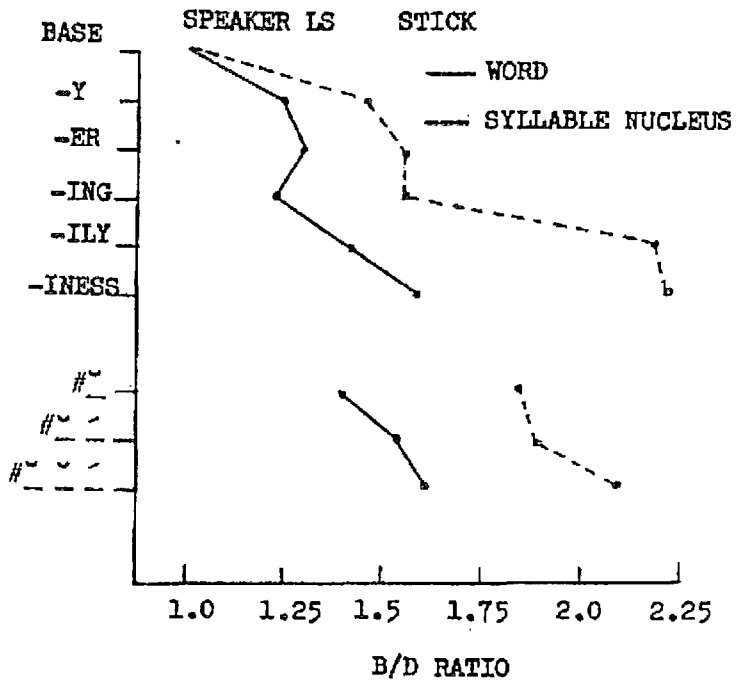
Table VI. Mean durations (in milliseconds), standard deviations, and B/D ratios for words derived with -<u>y</u> and -<u>ness</u>, in which the -<u>ness</u> words are derived by a two-cycle operation from the base.

Speaker R.G.				Speaker L.S.			
Utterance	Duration of base	σ	B/D ratio	Utterance	Duration of base	σ	B/D ratio
stick	401.55	29.45		stick	431.80	43.33	
stick-y	312.80	23.68	1.284	stick-y	346.00	34.44	1.248
sticky	513.25	37.52		sticky	557.45	36.59	
sticky-ness	376.75	17.66	1.362	sticky-ness	388.50	24.34	1.435
sleep	409.70	18.96		sleep	442.45	39.62	
sleep-y	336.70	19.70	1.217	sleep-y	363.40	19.64	1.388
sleepy	517.55	26.58		sleepy	544.20	30.99	
sleepy-ness	369.65	14.15	1.400	sleepy-ness	392.20	18.46	1.454
speed	511.50	34.95		speed	574.25	30.00	
speed-y	359.75	15.09	1.422	speedy	394.85	23.89	1.455
speedy	529.95	26.23		speedy	597.40	16.94	
speedy-ness	396.35	16.60	1.337	speedy-ness	410.55	31.19	1.413
shade	454.10	28.88		shade	454.65	35.84	
shade-y	327.20	20.08	1.388	shade-y	321.65	20.72	
shady	477.90	25.81		shady	490.60	24.43	
shady-ness	346.30	16.46	1.380	shady-ness	329.70	23.87	1.488

Table VII. Average B/D ratios (speakers R.G. and L.S. combined)

	stick, sleep		shade, speed	
	WORD	SN	WORD	SN
Base	1.00	1.00	1.00	1.00
-Y	1.242	1.425	1.420	1.669
-ER	1.262	1.476	1.425	1.706
-ING	1.256	1.474	1.469	1.799
-ILY	1.351	1.774	1.599	2.126
-INESS	1.492	1.931	1.716	2.415
# ´	1.388	1.638	1.409	1.873
# ˘ ´	1.518	1.857	1.626	2.194
# ˘ ˘ ´	1.497	1.914	1.683	2.502

Fig. 1. B/D ratios for the words <u>stick</u> and <u>sleep</u> and their syllable nuclei for speaker LS. The base word and the derivative forms are indicated on the vertical axis; the horizontal axis is calibrated for ratios of duration of base word/duration of the base part of the derived word.



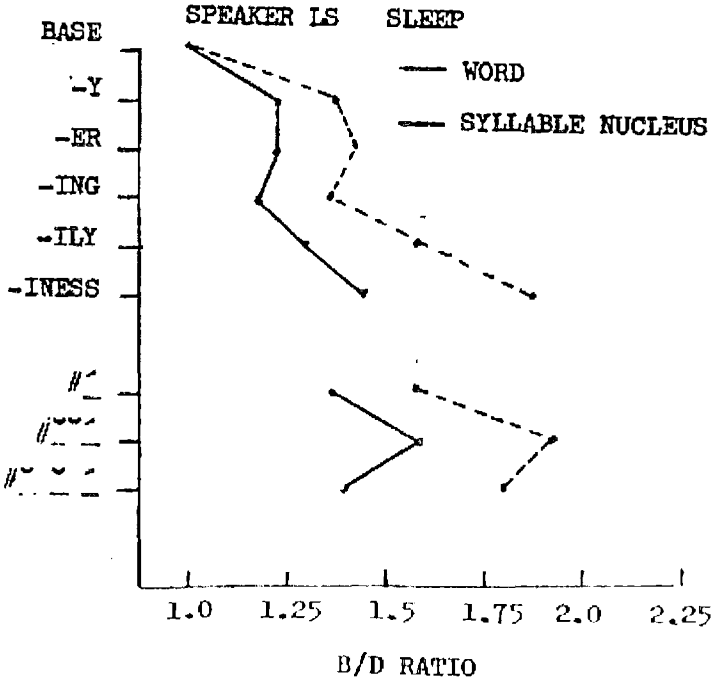
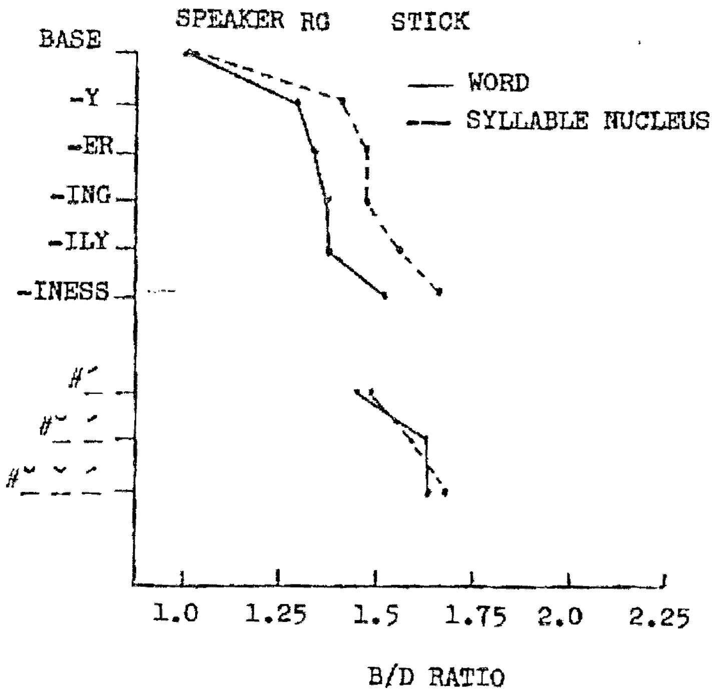


Fig. 2. B/D ratios for the words <u>stick</u> and <u>sleep</u> and their syllable nuclei for speaker RG. The base word and the derivative forms are indicated on the vertical axis; the horizontal axis is calibrated for ratios of duration of base word/duration of the base part of the derived word.



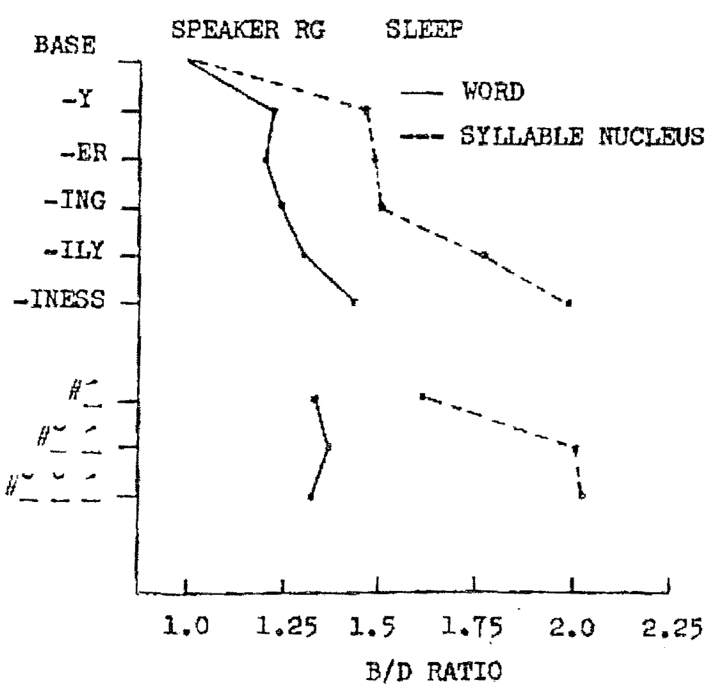
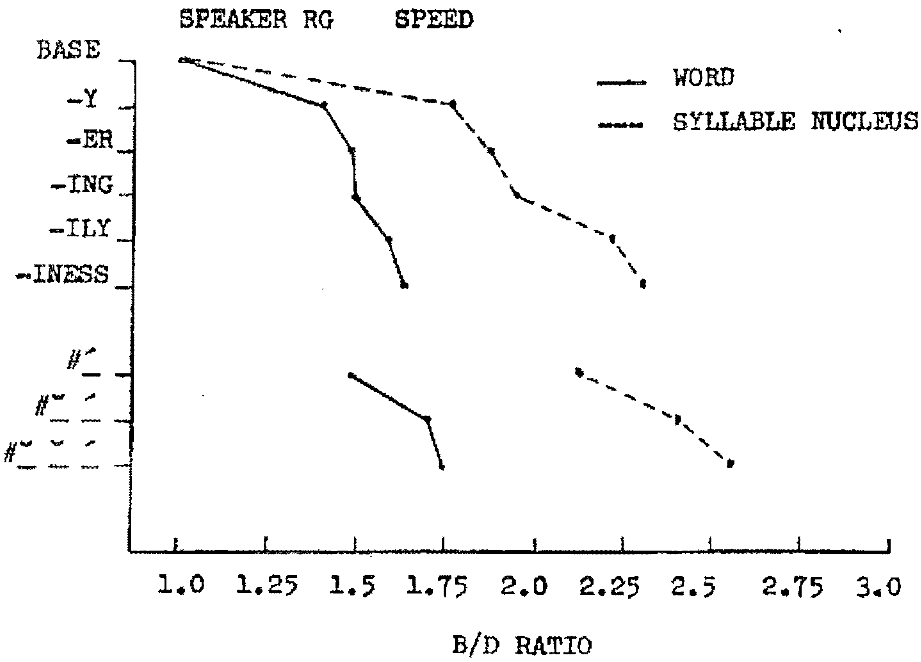


Fig. 3. B/D ratios for the words <u>speed</u> and <u>shade</u> and their syllable nuclei for speaker RG. The base word and the derivative forms are indicated in the vertical axis; the horizontal axis is calibrated for ratios of duration of base word/duration of the base part of the derived word.



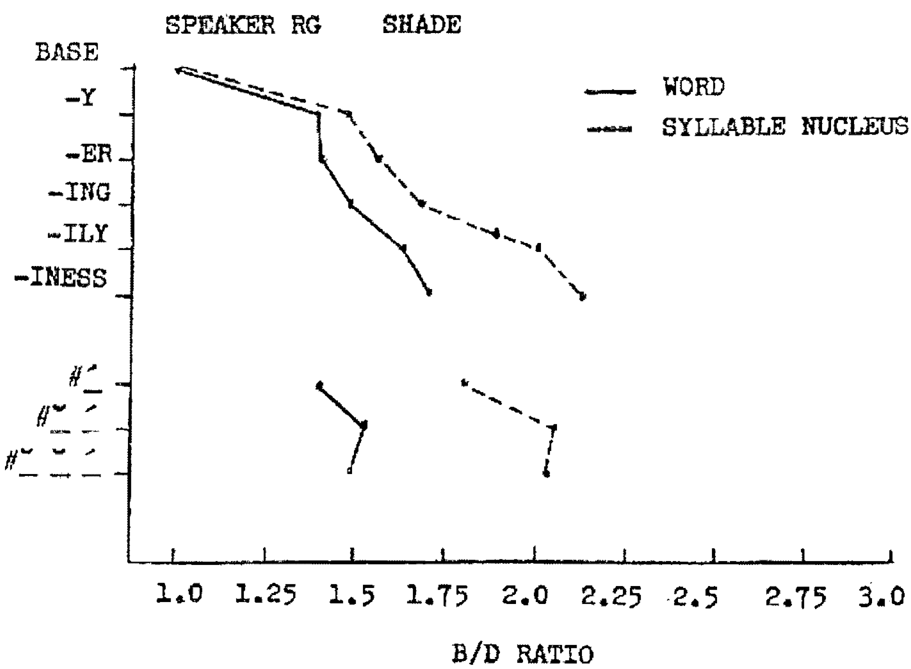


Fig. 4. B/D ratios for thw words speed and shade and their syllable nuclei for speaker LS. The base word and the derivative forms are indicated on the vertical axis; the horizontal axis calibrated for ratios of duration of base word/duration of the base part of the derived word.